RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _10/563, 655_

Source: _1FWP_

Date Processed by STIC: _1/18/06_

ENTERED

IFWP

RAW SEQUENCE LISTING   DATE: 01/18/2006
PATENT APPLICATION: US/10/563,655   TIME: 15:21:28

Input Set : A:\54-000410US.ST25.txt
Output Set: N:\CRF4\01182006\J563655.raw

```
   3 <110> APPLICANT: The Scripps Research Institute
   4        Anderson, J C
   5        Schultz, Peter G
   7 <120> TITLE OF INVENTION: COMPOSITIONS OF ORTHOGONAL LEUCYL-tRNA AND AMINOACYL-tRNA
   8        SYNTHETASE PAIRS AND USES THEREOF
  10 <130> FILE REFERENCE: 54-000410PC
C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/563,655
C--> 13 <141> CURRENT FILING DATE: 2006-01-05
  15 <160> NUMBER OF SEQ ID NOS: 72
  17 <170> SOFTWARE: PatentIn version 3.2
  19 <210> SEQ ID NO: 1
  20 <211> LENGTH: 88
  21 <212> TYPE: DNA
  22 <213> ORGANISM: artificial
  24 <220> FEATURE:
  25 <223> OTHER INFORMATION: mutant tRNA
  27 <400> SEQUENCE: 1
  28 gcgagggtag ccaagctcgg ccaacggcga cggactctag atccgttctc gtaggagttc     60
  30 gagggttcga atcccttccc tcgcacca                                        88
  33 <210> SEQ ID NO: 2
  34 <211> LENGTH: 88
  35 <212> TYPE: DNA
  36 <213> ORGANISM: artificial
  38 <220> FEATURE:
  39 <223> OTHER INFORMATION: mutant tRNA
  41 <400> SEQUENCE: 2
  42 gcgagggtag ccaagctcgg ccaacggcga cggactctaa atccgttctc gtaggagttc     60
  44 gagggttcga atcccttccc tcgcacca                                        88
  47 <210> SEQ ID NO: 3
  48 <211> LENGTH: 88
  49 <212> TYPE: DNA
  50 <213> ORGANISM: artificial
  52 <220> FEATURE:
  53 <223> OTHER INFORMATION: mutant tRNA
  55 <400> SEQUENCE: 3
  56 cccagggtag ccaagctcgg ccaacggcga cggactctaa atccgttctc gtaggagttc     60
  58 gagggttcga atcccttccc tgggacca                                        88
  61 <210> SEQ ID NO: 4
  62 <211> LENGTH: 89
  63 <212> TYPE: DNA
  64 <213> ORGANISM: artificial
  66 <220> FEATURE:
  67 <223> OTHER INFORMATION: mutant tRNA
```

P.6

RAW SEQUENCE LISTING   DATE: 01/18/2006
PATENT APPLICATION: US/10/563,655   TIME: 15:21:28

Input Set : A:\54-000410US.ST25.txt
Output Set: N:\CRF4\01182006\J563655.raw

```
69 <400> SEQUENCE: 4
70 gcgagggtag ccaagctcgg ccaacggcga cggacttcct aatccgttct cgtaggagtt    60
72 cgagggttcg aatcccttcc ctcgcacca                                      89
75 <210> SEQ ID NO: 5
76 <211> LENGTH: 89
77 <212> TYPE: DNA
78 <213> ORGANISM: artificial
80 <220> FEATURE:
81 <223> OTHER INFORMATION: mutant tRNA
83 <400> SEQUENCE: 5
84 gcgagggtag ccaagctcgg ccaacggcga cggacttcct aatccgttct cgtaggagtt    60
86 cgagggttcg aatccctccc ctcgcacca                                      89
89 <210> SEQ ID NO: 6
90 <211> LENGTH: 89
91 <212> TYPE: DNA
92 <213> ORGANISM: artificial
94 <220> FEATURE:
95 <223> OTHER INFORMATION: mutant tRNA
97 <400> SEQUENCE: 6
98 gcggggggttg ccgagcctgg ccaaaggcgc cggacttcct aatccggtcc cgtaggggtt    60
100 ccggggttca aatccccgcc cccgcacca                                     89
103 <210> SEQ ID NO: 7
104 <211> LENGTH: 88
105 <212> TYPE: DNA
106 <213> ORGANISM: artificial
108 <220> FEATURE:
109 <223> OTHER INFORMATION: mutant tRNA
111 <400> SEQUENCE: 7
112 gcggggggttg ccgagcctgg ccaaaggcgc cggacttcaa atccggtccc gtaggggttc    60
114 cggggttcaa atccccgccc ccgcacca                                      88
117 <210> SEQ ID NO: 8
118 <211> LENGTH: 77
119 <212> TYPE: DNA
120 <213> ORGANISM: artificial
122 <220> FEATURE:
123 <223> OTHER INFORMATION: mutant tRNA
125 <400> SEQUENCE: 8
126 ccggcggtag ttcagcaggg cagaacggcg gactctaaat ccgcatggcg ctggttcaaa    60
128 tccggcccgc cggacca                                                  77
131 <210> SEQ ID NO: 9
132 <211> LENGTH: 90
133 <212> TYPE: DNA
134 <213> ORGANISM: Escherichia coli
136 <400> SEQUENCE: 9
137 ggagagatgc cggagcggct gaacggaccg gtctctaaaa ccggagtagg ggcaactcta    60
139 ccgggggttc aaatcccccct ctctccgcca                                   90
142 <210> SEQ ID NO: 10
143 <211> LENGTH: 91
144 <212> TYPE: DNA
```

RAW SEQUENCE LISTING   DATE: 01/18/2006
PATENT APPLICATION: US/10/563,655   TIME: 15:21:28

Input Set : **A:\54-000410US.ST25.txt**
Output Set: **N:\CRF4\01182006\J563655.raw**

```
145 <213> ORGANISM: artificial
147 <220> FEATURE:
148 <223> OTHER INFORMATION: mutant tRNA
150 <400> SEQUENCE: 10
151 ggagagatgc cggagcggct gaacggaccg gtcttcctaa accggagtag gggcaactct    60
153 accgggggtt caaatccccc tctctccgcc a                                   91
156 <210> SEQ ID NO: 11
157 <211> LENGTH: 88
158 <212> TYPE: RNA
159 <213> ORGANISM: Halobacterium sp. NRC-1
161 <400> SEQUENCE: 11
162 gcgaggguag ccaagcucgg ccaacggcga cggacucaag auccguucuc guaggaguuc    60
164 gaggguucga aucccuuccc ucgcacca                                       88
167 <210> SEQ ID NO: 12
168 <211> LENGTH: 89
169 <212> TYPE: RNA
170 <213> ORGANISM: artificial
172 <220> FEATURE:
173 <223> OTHER INFORMATION: consensus tRNA
175 <400> SEQUENCE: 12
176 gcgggggttg ccgagccugg ccaaaggcgc cggacuuccu aaucccgucc cguaggggutt    60
178 cggggguuca aauccccgcc cccgcacca                                      89
181 <210> SEQ ID NO: 13
182 <211> LENGTH: 2799
183 <212> TYPE: DNA
184 <213> ORGANISM: Archaeoglobus fulgidus
186 <400> SEQUENCE: 13
187 atgagcgatt tcaggataat tgaggagaag tggcagaagg cgtgggagaa ggacagaatt    60
189 tttgagtccg atcctaatga gaaggagaag tttttttctca caattcccta tccttacctt   120
191 aatggaaatc ttcacgcagg tcacacgaga accttacaa ttggcgatgc cttcgccaga    180
193 tacatgagaa tgaagggcta caacgttctc tttccctcg gctttcatgt tacgggcacc    240
195 ccaatcattg gccttgcgga gctcatagcc aagagggacg agaggacgat agaggtttac    300
197 accaaatacc atgacgttcc gctggaggac ttgcttcagc tcacaactcc agagaaaatc    360
199 gttgagtact tctcaaggga ggcgctgcag gctttgaaga gcataggcta ctccattgac    420
201 tggaggaggg ttttcaccac aaccgatgaa gagtatcaga gattcatcga gtggcagtac    480
203 tggaagctca aggagcttgg cctgattgtg aagggcaccc accccgtcag atactgcccc    540
205 cacgaccaga tcctgttga agaccacgac cttctcgctg gggaggaggc aactattgtt    600
207 gaatttaccg ttataaagtt caggcttgaa gatggagacc tcatttttccc ctgtgcaact   660
209 ctccgtcccg aaaccgtgtt tggcgtcacg aacatctggg taaagccgac aacctacgta    720
211 attgccgagg tggatgggga aaagtggttt gtgagcaaag aggcttacga gaagctcacc    780
213 tacacggaga aaaaagtcag gctgctggag gaggttgatg cgtcgcagtt cttcggcaag    840
215 tacgtcatag tcccgctggt aaacagaaaa gtgccaattc tgcctgcaga gtttgttgac    900
217 accgacaacg caacaggagt tgtgatgagc gttcccgcac acgctccttt tgacctggct    960
219 gccattgagg acttgaagag agacgaggaa acgctggcga agtacggaat tgacaaaagc   1020
221 gttgtagaga gcataaagcc aatagttctg attaagacgg acattgaagg tgttcctgct   1080
223 gagaagctaa taagagagct tggagtgaag agccagaagg acaaggagct gctggataag   1140
225 gcaaccaaga ccctctacaa gaaggagtac cacacgggaa tcatgctgga caacacgatg   1200
227 aactatgctg gaatgaaagt ttctgaggcg aaggagagag ttcatgagga tttggttaag   1260
229 cttggcttgg gggatgtttt ctacgagttc agcgagaagc ccgtaatctg caggtgcgga   1320
```

```
231 acgaagtgcg ttgttaaggt tgttagggac cagtggttcc tgaactactc caacagagag   1380
233 tggaaggaga aggttctgaa tcaccttgaa aagatgcgaa tcatccccga ctactacaag   1440
235 gaggagttca ggaacaagat tgagtggctc agggacaagg cttgtgccag aaggaagggg   1500
237 cttggaacga gaattccgtg ggataaggag tggctcatcg agagcctttc agactcaaca   1560
239 atctacatgg cctactacat ccttgccaag tacatcaacg caggattgct caaggccgag   1620
241 aacatgactc ccgagttcct cgactacgtg ctgctgggca aaggtgaggt tgggaaagtt   1680
243 gcggaagctt caaaactcag cgtggagtta atccagcaga tcagggacga cttcgagtac   1740
245 tggtatcccg ttgacctaag aagcagtggc aaggacttgg ttgcaaacca cctgctcttc   1800
247 tacctcttcc accacgtcgc cattttcccg ccagataagt ggccgagggc aattgccgta   1860
249 aacggatacg tcagccttga gggcaagaag atgagcaaga gcaaagggcc cttgctaacg   1920
251 atgaagaggg cggtgcagca gtatggtgcg gatgtgacga ggctctacat cctccacgct   1980
253 gcagagtacg acagcgatgc ggactggaag agcagagagg ttgaagggct tgcaaaccac   2040
255 ctcaggaggt tctacaacct cgtgaaggag aactacctga aagaggtggg agagctaaca   2100
257 accctcgacc gctggcttgt gagcaggatg cagagggcaa taaaggaagt gagggaggct   2160
259 atggacaacc tgcagacgag gagggccgtg aatgccgcct tcttcgagct catgaacgac   2220
261 gtgagatggt atctgaggag aggaggtgag aacctcgcta taatactgga cgactggatc   2280
263 aagctcctcg ccccctttgc tccgcacatt tgcgaggagc tgtggcactt gaagcatgac   2340
265 agctacgtca gcctcgaaag ctacccagaa tacgacgaaa ccagggttga cgaggaggcg   2400
267 gagagaattg aggaatacct ccgaaacctt gttgaggaca ttcaggaaat caagaagttt   2460
269 gttagcgatg cgaaggaggt ttacattgct cccgccgaag actggaaggt taaggcagca   2520
271 aaggtcgttg ctgaaagcgg ggatgttggg gaggcgatga agcagcttat gcaggacgag   2580
273 gagcttagga agctcggcaa agaagtgtca aatttcgtca agaagatttt caaagacaga   2640
275 aagaagctga tgctagttaa ggagtgggaa gttctgcagc agaacctgaa atttattgag   2700
277 aatgagaccg gactgaaggt tattcttgat actcagagag ttcctgagga aagaggagg   2760
279 caggcagttc cgggcaagcc cgcgatttat gttgcttaa                           2799
282 <210> SEQ ID NO: 14
283 <211> LENGTH: 2814
284 <212> TYPE: DNA
285 <213> ORGANISM: Methanobacterium thermoautotrophicum
287 <400> SEQUENCE: 14
288 gtggatattg aaagaaaatg gcgtgataga tggagagatg ctggcatatt tcaggctgac     60
290 cctgatgaca gagaaaagat attcctcaca gtcgcttacc cctaccccag tggtgcgatg    120
292 cacataggac acgggaggac ctacactgtc cctgatgtct atgcacggtt caagaggatg    180
294 cagggctaca acgtcctgtt tcccatggcc tggcatgtca caggggcccc tgtcataggg    240
296 atagcgcgga ggattcagag gaaggatccc tggaccctca aaatctacag ggaggtccac    300
298 agggtccccg aggatgagct tgaacgtttc agtgaccctg agtacatagt tgaatacttc    360
300 agcagggaat accggtctgt tatggaggat atgggctact ccatcgactg gaggcgtgaa    420
302 ttcaaaacca cggatcccac ctacagcagg ttcatacagt ggcagataag gaagctgagg    480
304 gaccttggcc tcgtaaggaa gggcgcccat cctgttaagt actgccctga atgtgaaaac    540
306 cctgtgggtg accatgacct ccttgagggt gaggggttg ccataaacca gctcacactc    600
308 ctcaaattca aacttggaga ctcatacctg gtcgcagcca ccttcaggcc cgagacaatc    660
310 tatggggcca ccaacctctg gctgaaccct gatgaggatt atgtgagggt tgaaacaggt    720
312 ggtgaggagt ggataataag cagggctgcc gtggataatc tttcacacca gaaactggac    780
314 ctcaaggttt ccggtgacgt caaccccggg gacctgatag ggatgtgcgt ggagaatcct    840
316 gtgacgggcc aggaacaccc catactcccg gcttccttcg ttgaccctga atatgccaca    900
318 ggtgttgtgt tctctgtccc tgcacatgcc cctgcagact catagccct tgaggacctc    960
320 aggacagacc atgaactcct tgaaaggtac ggtcttgagg atgtggttgc tgatattgag   1020
322 cccgtgaatg tcatagcagt ggatggctac ggtgagttcc ggcggccga ggttatagag   1080
324 aaatttggtg tcagaaacca ggaggacccc cgccttgagg atgccaccgg ggagctatac   1140
```

```
326 aagatcgagc atgcgagggg tgttatgagc agccacatcc ctgtctatgg tggtatgaag 1200
328 gtctctgagg cccgtgaggt catcgctgat gaactgaagg accagggcct tgcagatgag 1260
330 atgtatgaat tcgctgagcg acctgttata tgccgctgcg gtggcaggtg cgttgtgagg 1320
332 gtcatggagg accagtggtt catgaagtac tctgatgacg cctggaagga cctcgcccac 1380
334 aggtgcctcg atggcatgaa gataataccc gaggaggtcc gggccaactt tgaatactac 1440
336 atcgactggc tcaatgactg ggcatgttca aggaggatag gccttggaac aaggctgccc 1500
338 tgggatgaga ggtggatcat cgaacccctc acagactcaa caatctacat ggcatattac 1560
340 accatcgcac accgcctcag ggagatggat gccggggaga tggacgatga gttctttgat 1620
342 gccatattcc tagatgattc aggaaccttt gaggatctca gggaggaatt ccggtactgg 1680
344 tacccccttg actggaggct ctctgcaaag gacctcatag gcaatcacct gacattccat 1740
346 atattccacc actcagccat attccctgag tcagggtggc cccgggggc tgtggtctttt 1800
348 ggtatgggcc ttcttgaggg caacaagatg tcatcctcca agggcaacgt catactcctg 1860
350 agggatgcca tcgagaagca cggtgcagac gtggtgcggc tcttcctcat gtcctcagca 1920
352 gagccatggc aggactttga ctggagggag agtgaggtca tcgggacccg caggaggatt 1980
354 gaatggttca gggaattcgg agagagggtc tcaggtatcc tggatggtag gccagtcctc 2040
356 agtgaggtta ctccagctga acctgaaagc ttcattggaa ggtggatgat gggtcagctg 2100
358 aaccagagga tacgtgaagc cacaagggcc cttgaatcat tccagacaag aaaggcagtt 2160
360 caggaggcac tctatctcct taaaaaggat gttgaccact accttaagcg tgttgagggt 2220
362 agagttgatg atgaggttaa atctgtcctt gcaaacgttc tgcacgcctg gataaggctc 2280
364 atggctccat tcatacccta cactgctgag gagatgtggg agaggtatgg tggtgagggt 2340
366 tttgtagcag aagctccatg gcctgacttc tcagatgatg cagagagcag ggatgtgcag 2400
368 gttgcagagg agatggtcca gaataccgtt agagacattc aggaaatcat gaagatcctt 2460
370 ggatccaccc cggagagggt ccacatatac acctcaccaa aatggaaatg ggatgtgcta 2520
372 agggtcgcag cagaggtagg aaaactagat atgggctcca taatgggaag ggtttcagct 2580
374 gagggcatcc atgataacat gaaggaggtt gctgaatttg taaggaggat catcagggac 2640
376 cttggtaaat cagaggttac ggtgatagac gagtacagcg tactcatgga tgcatctgat 2700
378 tacattgaat cagaggttgg agccagggtt gtgatacaca gcaaaccaga ctatgaccct 2760
380 gaaaacaagg ctgtgaatgc cgttcccctg aagccagcca tataccttga atga 2814
383 <210> SEQ ID NO: 15
384 <211> LENGTH: 932
385 <212> TYPE: PRT
386 <213> ORGANISM: Archaeoglobus fulgidus
388 <400> SEQUENCE: 15
390 Met Ser Asp Phe Arg Ile Ile Glu Glu Lys Trp Gln Lys Ala Trp Glu
391 1               5                   10                  15
394 Lys Asp Arg Ile Phe Glu Ser Asp Pro Asn Glu Lys Glu Lys Phe Phe
395             20                  25                  30
398 Leu Thr Ile Pro Tyr Pro Tyr Leu Asn Gly Asn Leu His Ala Gly His
399         35                  40                  45
402 Thr Arg Thr Phe Thr Ile Gly Asp Ala Phe Ala Arg Tyr Met Arg Met
403     50                  55                  60
406 Lys Gly Tyr Asn Val Leu Phe Pro Leu Gly Phe His Val Thr Gly Thr
407 65                  70                  75                  80
410 Pro Ile Ile Gly Leu Ala Glu Leu Ile Ala Lys Arg Asp Glu Arg Thr
411                 85                  90                  95
414 Ile Glu Val Tyr Thr Lys Tyr His Asp Val Pro Leu Glu Asp Leu Leu
415             100                 105                 110
418 Gln Leu Thr Thr Pro Glu Lys Ile Val Glu Tyr Phe Ser Arg Glu Ala
419         115                 120                 125
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 01/18/2006
PATENT APPLICATION: US/10/563,655   TIME: 15:21:29

Input Set : A:\54-000410US.ST25.txt
Output Set: N:\CRF4\01182006\J563655.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:17; N Pos. 3750,3770
Seq#:18; N Pos. 3749,3769
Seq#:19; N Pos. 3749,3769
Seq#:20; N Pos. 3274,3294
Seq#:72; N Pos. 31,45,77

Invalid <213> Response:
Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:1,2,3,4,5,6,7,8,10,12,17,18,19,20,21,23,24,25,26,27,28,29,30,31,32,33,34
Seq#:35,36,72

VERIFICATION SUMMARY   DATE: 01/18/2006
PATENT APPLICATION: US/10/563,655   TIME: 15:21:29

Input Set : A:\54-000410US.ST25.txt
Output Set: N:\CRF4\01182006\J563655.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application Number
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:1013 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17 after pos.:3720
L:1196 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18 after pos.:3720
L:1379 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:3720
L:1546 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20 after pos.:3240
L:2260 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:72 after pos.:0
M:341 Repeated in SeqNo=72